Visitor Flow Management using Human-Robot Interaction at Expo.02

B.Jensen, G.Froidevaux, X.Greppin, A.Lorotte, L.Mayor, M.Meisser, G.Ramel, R.Siegwart

Autonomous Systems Lab - Swiss Federal Institute of Technology - CH-1015 Lausanne

Abstract

In this paper we will regard the task of operating a public mass exposition with several autonomous robots at a time. This implies questions regarding human-robot interaction, multi-robot control and interaction management. To enable human-robot interaction while guiding a tour we outline the SOUL environment. Multi-robot and interaction management are regarded with respect to visitor density and visitor flow. Concluding we will present and discuss results from the Swiss national exhibition Expo.02 in the time from 15.05.02 to 17.07.02, corresponding to 5293 hours of total robot operation time up to date and in interaction with 283319 visitors.

1. Introduction

Public space experiences in recent years are proof of a remarkable progress in mobile robotics. This enabled the operation of a public mass exposition with ten autonomous mobile robots at a time during the Swiss national exhibition Expo.02.
Having several identical robots serving as tour-guide and main attraction of an exposition during a five-month period from 15.05.02 to 20.10.02 created a special situation. Men and machine operating in the same space make reliable and safe robot operation is mandatory. Ten and a half-hours operation per day, seven days per week over the exposition period imposed high demands on robotics hardware. In addition to this, visitor flow and fun factor of an exposition are important to operators and financiers of a public mass exposition.
To meet these requirements, the interactive mobile tour-guide RoboX has been recently developed by our lab. Developing the interactive part for the exposition meant always taking into account the demand for visitor flow and entertainment. These criteria translate more or less directly into guided tour and unconstrained interaction. Our solution is embodied in the SOUL (Scenario Object Utility Language) system [8] controlling guided tour and interaction together.
The fact of having several robots at disposal makes them easier and faster available for the visitors, but requires a resource management for the exposition space. The autonomous nature of our robot evokes the question of centralized or distributed system architecture, which we will regard later on.

Figure 1: RoboX interacting with people visiting Expo.02.

Closely related with the multi-robot control, we try to support the natural visitor flow direction from entry to exit by constraining the displacement of the robot.
In general high visitor density and a rapid visitor flow constrain interaction. Since these parameters are external, we seek a system allowing for a maximum of interaction under the current conditions.
Concluding we will evaluate these elements under real world conditions based on experience gained at the Expo.02.

2. Related work

We will look at mobile robot experiences in public spaces, arguing that the mobility of the platform and the direct presence of both human and robot render interaction particularly interesting. We find the importance of improving human robot interfaces [1], to help visitors in interacting with mobile robots. Face and emotional state machines were found useful elements for tour-guide-robots [2]. The Mobot Museum Robot Series [3,4] focused on the interaction. Robustness and reliability was identified as an important part of a public robot. Several experiences with the museum robots showed further that the visitors do not always behave cooperatively with the robot and switch between seeing it as a simple machine or a tour-guide. Another permanent installation is at the "Deutsches Museum für

Kommunikation” in Berlin, where three robots welcome the visitors and invite them to play with a ball [5].
Summarizing, we can state that the development of public robots has to take into account the differences in visitors’ behavior. First of all, the robot needs to sense the presence of visitors in order to react appropriately. We may distinguish if the robot is seeking an interaction or if it is already giving a tour and interacting with someone else [4]. It was further found that the time visitors spend with the robot is not easily predictable or controllable. Some visitors get bored after a couple of minutes with the robot, others spent days with it. During this time the visitors’ behavior changed from collaborative to investigative interaction.

3. RoboX

During Expo.02, the time which visitors can spend with RoboX is rather limited. We decided to use intuitive means of communication in order to use this time as efficiently as possible. The design of the robot should use common features for communication, situating its appearance somewhere between anthropomorphic and machine. The face of RoboX is intended the source of communication helping the visitors to feel more comfortable when communicating with the robot.

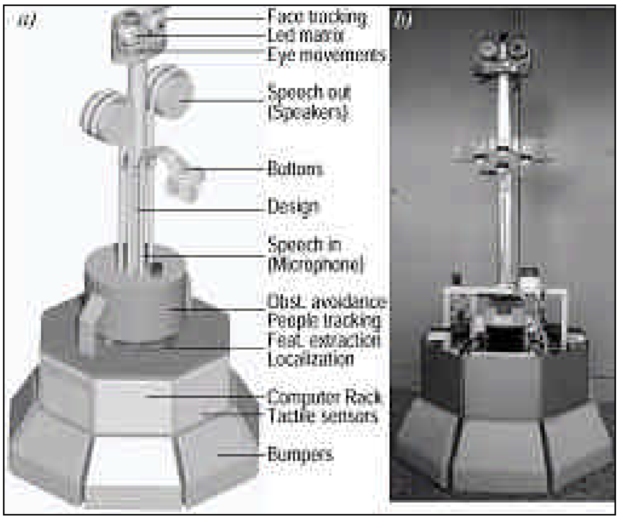

Figure 2: Outline of RoboX elements and photo of the first prototype.

Even though collaborative interaction will mainly take place between one visitor and the robot, we anticipate that a certain audience of other visitors will follow this interaction. For good visibility we constructed RoboX (figure 2) to be of approximately average visitor’s height. Basically, the robot consists of a mobile base with an interactive top, making the face easy to look at.
Two differentially driven wheels located at the center of the robot allow on the spot turns. Two castor wheels, one at its back and one, with a suspension at its front, ensure the stability of the mobile base. Obstacle avoidance and reliable localization [6] ensure that the robot knows at all times its position and does not collide with visitors or parts of the exposition.
As an additional means of security, touch sensitive plates and foam bumpers ensure that the robot stops if running into anything. Two SICK Laser scanners mounted at knee height provide environmental information for navigation and interaction. A camera mounted in one of the robot’s eyes provides additional information for the interaction.
Furthermore, the mobile base houses motor controllers, batteries for 10h autonomy, a PowerPC 750 clocked at 400 MHz dedicated for navigation and obstacle avoidance and a Pentium III running at 700 MHz, 128 MB RAM on Windows 2000 for all interaction tasks. Both computers can communicate with each other over a 10 Mbit/sec local Ethernet and with a central computer over wireless interfaces to allow monitoring the state of the robot for security reasons. Technical details are discussed in [7].

4. Interaction at Expo.02

Interaction of visitors with several robots in a public exposition is a complex task. First of all we will present how interaction between RoboX and a visitor is realized. We will distinguish static and dynamic elements, which help in making each tour of the robot individual. By taking into account dynamic elements, which we will precise later on we aim at giving the robot an aura conscious of its environment.
Since RoboX is giving a tour it will stop at several stations and supply information related to a certain part of the exposition. With the several RoboXs running at the same time we faced the problem of multi-robot coordination to avoid having several robot s intending to go at the same place at the same time.
Finally we will present how parameters like visitor flow and visitor density are taken into account to provide the most of interaction under the current conditions of the exposition.

4.1 SOUL

We will briefly present SOUL, controlling interaction on RoboX. It aims at combining elements of a guided tour with human-robot interaction. The tour the robot is giving presents a certain amount of information on several parts of the exposition. They will change rarely if ever, for the period of the exposition. Henceforth static scenarios can easily represent this information. A scenario is in the SOUL context the succession of robot actions as speaking, moving and similar actions for a limited amount of time.
Intelligent appearance can hardly be achieved by repeating these scenarios over and over again. Therefore we use methods of changing presentation and methods of adaptive behavior to avoid repetition.

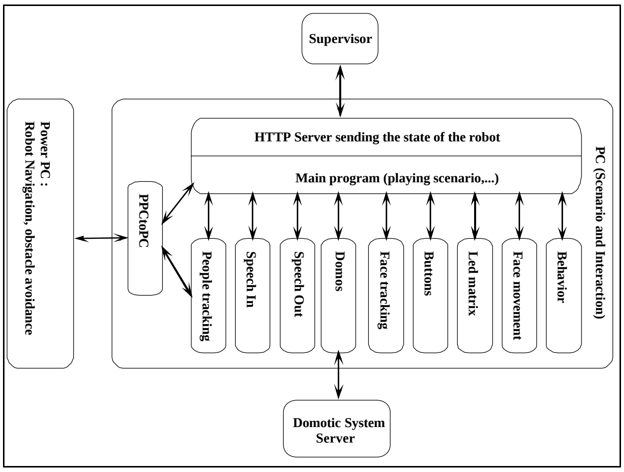

Figure 3: Structure of the interactive system. The supervisor is a separate computer allowing the operator to monitor of the robot's operation.

One way to avoid repetitive behavior is to provide several alternatives of the text and actions presented. Thus changing the method of presentation. The tools available to the SOUL system for creating such scenarios are exhibited in figure 3.

In addition to this permutation approach, we aimed at having a robot responding to a couple of dynamic events, which can occur during a tour. This changes its behavior. Such events can be visitors are blocking the robot or even hitting its bumpers. They are playing with the buttons without being asked to or are pressing the emergency button. The battery of the robot is low or other. From the point of view of interaction one can see these signals as a certain acceptance of the robot by the visitor. From the point of view of a guided tour, however they are exceptions and are treated by SOUL as such. Technically SOUL will interrupt the current scenario and execute a corresponding exception scenario telling the visitor that it is aware of his actions, before resuming the tour. RoboX will treat one exception at a time.

4.1.1 SOUL sensors

RoboX is using several sensors and algorithms to achieve awareness of its environment. Simple switches detect events like visitors pressing the emergency button, the interactive buttons or hitting the bumpers. The obstacle avoidance provides information when visitors are blocking the robot.

In addition the robot is aware of visitor presence in its surrounding by means of face and motion tracking [8].

4.1.2 SOUL expression

There are three interfaces available to communicate with the visitor. To express itself RoboX is using synthesized speech in English, French, German and Italian using Mbrola [9] and LAIPTTS [10].

Figure 4: Three facial expressions. From left to right: happy, surprised, and angry.

The interactive buttons can be illuminated to indicate in which mode they are in (language choice, yes/no, etc.).

For visitors the most expressive part remains the face (figure 4) imitating several grimaces and by means of a small LED display mounted in one of the eyes display symbols and short animations.

4.1.3 Behavior component

Our aim was to create individual tours according to the visitor's action, until yet their action affected the tour only shortly by starting the appropriate exception scenario. With the behavior component presented in [11] RoboX started to accumulate impressions during a tour and to adapt its behavior accordingly.

Here we have to distinguish two main cases in which RoboX uses the expressions. The first case happens to emphasize or illustrate its speech and is controlled directly by the scenario running. In the second case, the expressions are more like the mirror of the subject's emotions. For the representation of this internal state we chose the Arousal-Valence-Stance affect space [12], because of its three dimensional representation which is very intuitive to use. The robot current state is therefore defined as a point in the three-dimensional AVS space (see figure 5). In this space, six basic expressions regions are defined as: sadness, disgust, joy, anger, surprise and fear.

Also, we use the origin of the space as a reference expression that can be considered as a calm state. Of course, other expression regions can be defined in this space. But, we decided to limit ourselves to those seven regions in order not to overwhelm the visitor with many reactions to subtle for our robot expressive capacities.

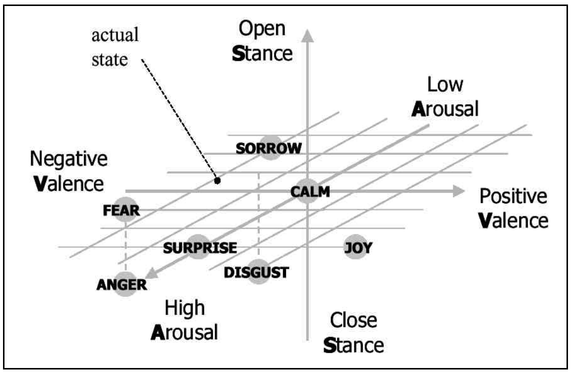

Figure 5: Representation of the six basic expressions and the neutral expression in the AVS space.

	Pitch	Rate	Volume
Fear	high	very fast	medium
Surprise	very high	very fast	very loud
Joy	high	fast	loud
Sorrow	little low	slow	very soft
Disgust	low	very slow	soft
Anger	very low	very slow	very loud

Figure 6: Variation of pitch, rate and volume for the standard expressions.

The internal state is mainly communicated using the synthesized voice, face movements in some cases symbols are shown on the LED screen. Figure 6 shows how the internal state effects the synthesized voice.

4.2 Multi-robot coordination

Figure 7 exhibits the layout of the exposition. Presentation stations are defined near particular objects in the expositions. There are several places where robots welcome visitors, thus tours can start simultaneously. At the time of writing there are fifteen presentation stations all over the exposition space. Finally there are goodbye stations close to the exit.

Each station corresponds to one scenario in the SOUL system, providing visitors with the necessary explanatory or entertaining information. Tours can be created by a succession of several presentation stations. Two stations are except from the tours and are permanently occupied with a dedicated robot. They have tasks of taking pictures from the visitors and presenting a slide show. In these cases the tour consists of one station only.

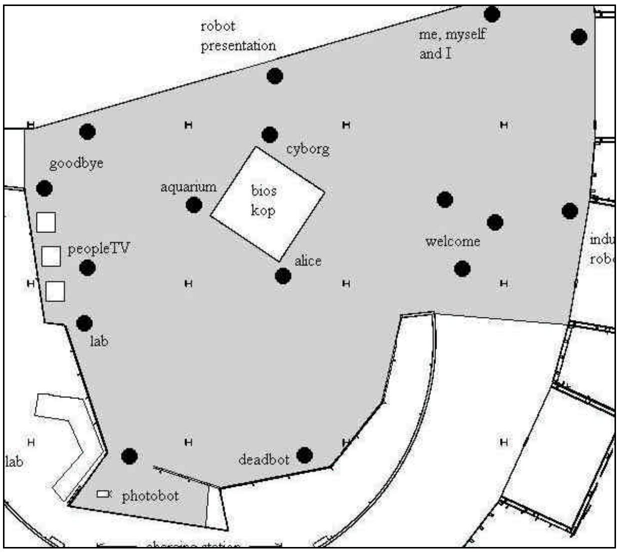

Figure 7: Scheme of the 315 m^2 exposition area with the presentations stations shown.

Working with multiple robots makes resource allocation an important point. In order to avoid having several robots presenting the same object an assignment has to be made at a certain moment.

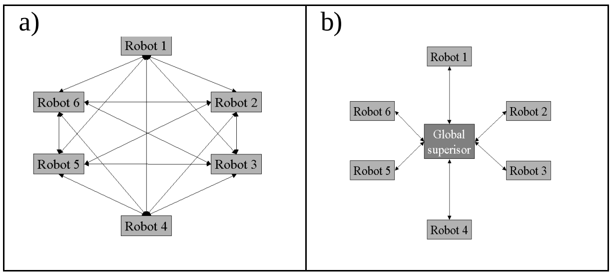

Figure 8: Communication structure a) without central server, b) with central unit.

In the beginning we solved this problem by assigning several stations exclusively to one tour which was operated by one robot all day. The tours were designed to have robots working spatially separated in order to avoid collisions among robots.

With ten robots operating the exposition this was no longer feasible, since it would result in tours of one or two stations only and thus quasi-static mobile robots. Improved obstacle avoidance allowed the robots to see each other and to avoid collisions. This enabled a dynamic assignment of stations to a robot for the duration of its presentation. The station is released thereafter and can be used by other robots.

This is modeled by a list of all stations and their state. Stations are free until reserved by a robot. The robot can chose among the free stations in order to avoid deadlocks. Care has to be taken that robots decide successively to avoid several robots choosing the same goal.

Figure 8 shows two different communication architectures for the assignment process. On the left side communication takes place among the robots only. Even though this uses only intelligence and information present in the robots it requires a complex communication. Each robot has to communicate with all other robots and needs to monitor which robots are currently active. Assuming N robots at hand all reserving one station this results in $N \cdot (N-1)$ communications.

By adding a central instance as shown in figure 8 b) this number drops to N communications. We opted for this solution since it results in a much easier and thus more reliable communication scheme. Technically this global instance could be run on one selected robot, so that the group of robots still can be considered as an autonomous system.

Multi-robot coordination in our case is based on local decisions by each robot. When terminating a presentation the robot will ask the state of all exposition stations from the global instance. This request blocks the global supervisor until the robot reserves a specific station. The decision, which station to reserve is based on the free stations, the list of stations included in this tour and the stations already visited. The first free and unvisited station in the tour list is reserved.

4.3 Visitor density and interaction

Expo.02 was considered a mass exposition with several thousands visitors per day. During the preparation of this project we anticipated up to 500 visitors per hour, which assuming a 15 minutes stay inside the exposition results in 125 visitors which are at the same time enjoying the robots.
Visitor behavior can hardly be anticipated. To ensure a functioning of the exposition even with a lot of visitors on the hand and to provide intensive interaction when viewer visitors are in the exposition, four exposition modes were defined:

1. *Wait for visitor*: with few visitors, so that robots wait for one to come close enough before starting to talk and ask him which station he would like to see.
2. *Visitor's choice*: more visitors, so that the robot can ask permanently whether the visitor wants to go to a station without talking to no one.
3. *Robot's ch*oice: even more visitors, so that the robot will decide what is the next station and go there without asking.
4. *No move*: too many visitors for the robot to move, so that each robot will stay with one station and present it permanently.

The exposition mode is defined manually by the staff. It is included in the data provided by the global supervisor, so every time the robot requests the state of the exposition stations it receives an update of the state and can adapt accordingly. Figure 9 shows how this is taken into account by the SOUL system:

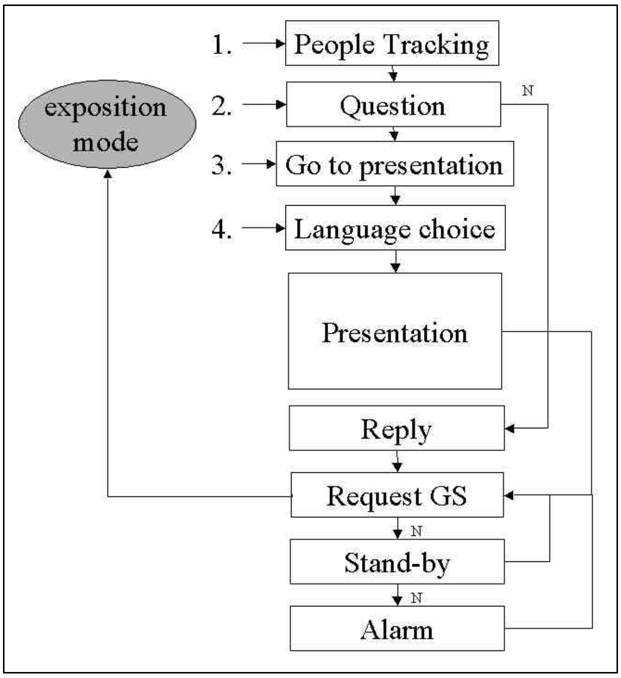

Figure 9: Structure of the SOUL sequence for a typical presentation station.

Depending on the exposition mode the scenario starts either with people tracking (*wait for visitor*), the question "Do you want to see … ?" (*visitor's choice*), the robot moving to the station (*robot's choice*) or directly with the language choice (*no move*).
These blocks are executed successively except if the visitor declines to go to a station. In this case SOUL jumps directly to the reply block commenting in some way the visitor's decision.
The request from the global supervisor is executed either after the reply block or after the presentation of a station. It provides all empty stations at this time the choice is made as explained in the paragraph above.
If no empty station is available and all empty station already have been visited during this tour the robot can not go on. Then it starts one of several stand-by scenarios. These are presentations, which are not located at a specified place in the exposition. The robot talks about itself, sings or makes funny faces. Thus the robot gains time during which a presentation station may be released by another robot.
After the stand-by scenario the robot request once again the exposition state to find a free presentation station. If one is found the next scenario is run. Otherwise the robot continues to play stand-by scenarios and to request the global supervisor until either a presentation station is available or it has run out of stand-by scenarios. In the latter case the global supervisor will give an alarm and the staff can interact. Starting the robot with another tour may solve this problem.
To avoid having several robots giving the same presentation a station remains blocked by one robot until it starts moving on to the next station.

4.4 Visitor flow

We estimated the average visit to 15 minutes in order to meet the visitor flow requirements. Previous test in our lab [8] proved it difficult for the robot to make visitors leave. In general their interest span is not directly related to the duration of a tour.
Visitor flow is channeled by two factors. First of all the number of stations the robot visits. The robot to visits S stations before it executes the goodbye scenario, which is located near the exit. By this proximity we aim at encouraging visitors to leave. The goodbye scenario is special in the way that it resets the list of stations visited during a tour and sets the counter of stations visited back to zero.
Throughout the exposition a tour will always lead visitors closer to the exit. This eases navigation and helps maintaining the visitor flow. Technically this is realized by a list of possible next presentation stations. Each presentation scenario is assigned an individual list, containing only stations to support the direction of the main visitor flow. When requesting exposition state from the global supervisor the robot will seek only stations which it has not yet visited and are closer to the exit than it is currently.

5. Results

In the period from 15.05.02 to 17.07.02 an average number of 4427 people were visiting the exposition every day. The minimal number of visitors was one time 2299 the maximum achieved was 5473. The average number results in a visitor flow of 422 persons per hour on 315 m^2 exposition space with up to ten robots in operation. This corresponds to a load of 84.3% percent of the planned maximal flow of 500 visitors. The maximum flow corresponds to a load of 104%.

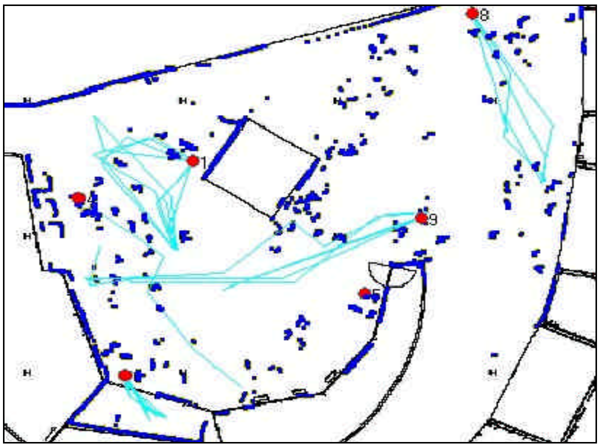

Figure 10: Map of exposition with 6 robots and laser scanner data showing visitors and robots (circles).

The global supervisor system is operational since the 01.07.02. Until yet the exposition mode *visitor's choice* was active approximately 95% the mode *robot's choice* 5% of the time. We experienced ten days with more than 5000 visitors, even in this crowded environment robots managed to move to their goal in a reasonable time, so that the mode *no move* was never used. Up to date the mode *wait for visitor* was never used, since the robots are most of the times surrounded by interested visitors anyway. Figure 10 shows a typical situation.

With currently three stand-by scenarios, alarms of a robot running out of those scenarios occurred approximately once a week. With two additional stand-by scenarios we aim at reducing this rate further.

Visitors stay between 10 and 45 minutes with the robots. We tried to control this by changing the tour length from two to ten stations without noticing an impact an the visitor's stay. People just move on to the next robot or even stay with the current one. Here enhanced environmental information, like motion information of the visitor or face recognition might help creating more convincing scenarios. We found that visitors quit a robot approximately after four stations, which is the actual tour length. The average number of visitors during the 17 days of operation of the global supervisor rose slightly to 4576 per day. This makes it hard to prove a quantitative effect on the visitor flow. However, observation of the crowd shows that visitor appreciated having the choice to go to a station. This adds a little interactive element to the tour.

6. Conclusion

During over 5293 hours of operation, 283319 visitors interacted with the robots in the time from 15.05.02 to 17.07.02. SOUL seems to provide an appealing compromise of a guided tour and unconstrained interaction. For the last two and a half weeks the exposition was running with a multi-robot resource control scheme taking into account the visitor density and supporting visitor flow.

Quantitative parameters like visitor flow and density meet the planning parameters. By enhancing environmental perception aim at creating even more convincing human-robot interaction.

References

[1] Burgard W., A. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, S. Thrun, “Experiences with an interactive museum tour-guide robot”. AI (114), 1999, pp. 3-55.

[2] Thrun S., M. Bennewitz, W. Burgard, A. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, D. Schulz, “MINERVA: A Second-Generation Museum Tour-Guide Robot”. In: Proc. ICRA, USA; 1999, vol.3; pp.1999-2005.

[3] Willeke T., C. Kunz, I. Nourbakhsh, “The History of the Mobot Museum Robot Series: An Evolutionary study”. In: Proc. FLAIRS 2001, Florida.

[4] Nourbakhsh I., J. Bobenage, S. Grange, R. Lutz, R. Meyer, A. Soto, "An Affective Mobile Educator with a Full-time Job". AI, 114(1-2), October 1999, pp. 95-124.

[5] Graf, B., W. Baum; A. Traub, R. Schraft, "Konzeption dreier Roboter zur Unterhaltung der Besucher eines Museums". In: VDI-Berichte 1552, VDI-Verlag, Düsseldorf, 2000, pp. 529-536.

[6] Arras, K., R. Philippsen, M. Battista, M. Schilt, R. Siegwart “A Navigation Framework for Multiple Mobile Robots and its Application at the Expo.02 Exhibition”. Workshop: Robots in Exhibitions, IEEE/RSJ.IROS, 2002, Switzerland.

[7] Tomatis N., G. Terrien, R. Piguet, D. Burnier, S. Bouabdhalla, R. Siegwart, “Design and System Integration for the Expo.02 Robot”. Workshop: Robots in Exhibitions, IEEE/RSJ.IROS, 2002, Switzerland.

[8] Jensen, B., G. Froidevaux, X. Greppin, A. Lorotte, L. Mayor, M. Meisser, G. Ramel, “The interactive autonomous mobile system RoboX”. In: Proc. IEEE/RSJ.IROS, 2002, Switzerland.

[9] Dutoit, T., et al, "The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes". In: Proc. ICSLP'96.

[10] Siebenhaar-Rölli, B. et al, “Phonetic and Timing Considerations in a Swiss High German TTS System”. In: E. Keller, G. Bailly, A. Monaghan, J. Terken & M. Huckvale, “Improvements in Speech Synthesis”, pp. 165-175, Wiley & Sons, 2001.

[11] Mayor L., B. Jensen, A. Lorotte, R. Siegwart, “Improving the expressiveness of mobile robots”. In: Proc. ROMAN(2002), Berlin, Germany.

[12] Ekman P., R. Davidson, “The Nature of Emotion: Fundamental Questions”, Oxford University Press, New York, 1994.